DOCUMENT RESUME

ED 170 358 TM 008 782

AUTHOR Williams, John D.; Wali, Mohan K.
TITLE Missing Cells and a Curious Case of Degrees of Freedom.
INSTITUTION North Dakota Univ., Grand Forks. Center for Teaching and Learning.
SPONS AGENCY Bureau of Mines (Dept. of Interior), Washington, D.C.
NOTE 14p.; Paper presented at the Annual Meeting of the American Educational Research Association (63rd, San Francisco, California, April 8-12, 1979)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Analysis of Variance; Control Groups; Experimental Groups; Hypothesis Testing; Mathematical Models; *Matrices; *Research Design; *Sampling; Statistical Studies
IDENTIFIERS *Incomplete Block Designs

ABSTRACT

A solution is proposed for analysis of variance procedures with missing cells, such as may occur when a control group is not assigned to any of the rows or columns of the various experimental groups. Mathematical models for two-way design are presented which define several variables; as well as row effect, column effect, and row and column effects. Full rank models, also recommended for disproportional data, are also presented. Restricted models are used to test hypotheses; this is also illustrated after dropping a vector. Two examples are described in which this research design may have educational applications: (1) the experimenter has a two-way treatment classification procedure, but the true control group is administered neither of the treatment combinations; or (2) cultural bias may be accounted for when a test is more fair in particular geographic areas; distances from these areas are represented in diagrams. (Author/GDC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

1979 AERA
Session 16.30

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED170354

MISSING CELLS AND A CURIOUS CASE OF DEGREES OF FREEDOM[1]

John D. Williams and Mohan K. Wali
University of North Dakota

PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
John D. Williams
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ABSTRACT

An experimental sampling procedure for plant communities on surface mined areas yielded missing cells and caused a further problem of yielding a "total" number of degrees of freedom equal to N rather than the usual N-1. The discrepancy occurred because the degrees of freedom are not necessarily additive for all missing cell designs. A solution which may circumvent this problem is proposed.

While many researchers in applied statistics have recently considered the two (or n) way disproportional fixed effects design, the missing cells case has not been fully reported. Because missing cells do occur either by design, distribution of natural populations, or topographic influences (as in the present case), they deserve due consideration. Perhaps the most extensive treatment of the cases of missing cells has been discussed by Searle (1971).

Missing cells in our study (Wali 1977) occurred in a systematic sampling of plant communities at six sites: five were previously surface mined for coal and one site in an unmined prairie was chosen for comparison. Prior to state and federal regulations on reclamation, these surface mined areas looked like a series of cone-like mounds (commonly referred to as spoil banks). Since topography, slope angle and aspect are important determinants of plant distribution, it was decided to obtain quantitative information on the distribution of plant species from crest, upper slope and mid-slope positions along eight cardinal compass directions (Fig. 1). In the sampling procedure, species composition of each sub-area was determined at every major site. This yielded (Fig. 1) 8 sampling areas in each of the upper and lower slope positions and 1 from the crest for a total of 17. The criterion measures were the types and numbers of plant species in each sample. The two-way classification with the number in each cell is shown in Table 1.

[1] Contribution from "A Systems Approach to Reclamation of Strip-Mined Areas in North Dakota" funded by U.S.D.I., Bureau of Mines, Grant No. 0264001 to M.K.W.

TM008 282

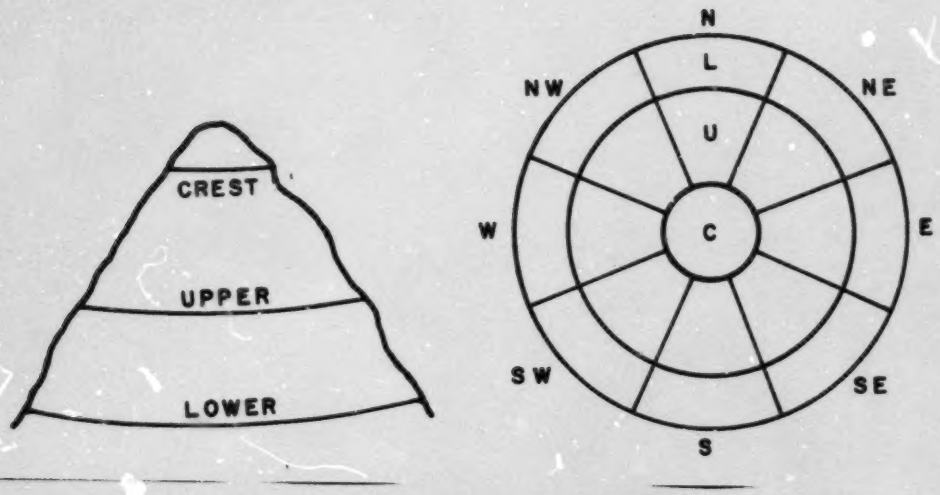

Fig. 1. (a) a diagrammatic representation of a "spoil bank" showing the position of the crest, and the upper and lower slopes; (b) a top view of the same showing sub-areas sampled along compass directions.

Table 1

Number of Samples from Five Strip-Mined and One Unmined Area by Sampling Units

	Direction								
	Crest	N	NE	E	SE	S	SW	W	NW
LEVEL CREST	12	-	-	-	-	-	-	-	-
UPPER	-	12	12	12	12	12	12	12	12
LOWER	-	12	12	12	12	12	12	12	12

NOTE: N = North octant; NE = Northeast octant; E = East octant; SE = Southeast octant; S = South octant; SW = Southwest octant; W = West octant; and NW = Northwest octant.

The unique qualitites of the research design can be seen by examining Table 1. When confronted by data gathered from such a design some might suggest that if the cell in the upper left-hand corner is dropped, there is no problem. However, such a suggestion begs the issue. Also, the actual three-way design is such that the proportionality holds for the other two ways of classification. An additional point is revealed by Table 1; the upper left hand cell does not enter in any way into the interaction. Thus, the interaction sum of squares and degrees of freedom (df = 7) would be the same whether or not the crest data are included.

Models for Two Way Design

First, several variables are defined

Y = the criterion score (number of distinct species in the subsample)
X_1 = 1 if the score is from a crest sample, 0 otherwise;
X_2 = 1 if the score is from an upper sample, 0 otherwise;
X_3 = 1 if the score is from the N octant; 0 otherwise;
X_4 = 1 if the score is from the NE octant; 0 otherwise;
X_5 = 1 if the score is from the E octant; 0 otherwise;
X_6 = 1 if the score is from the SE octant, 0 otherwise;
X_7 = 1 if the score is from the S octant, 0 otherwise;
X_8 = 1 if the score is from the SW octant, 0 otherwise;
X_9 = 1 if the score is from the W octant, 0 otherwise;
$X_{10} = X_2 \cdot X_3$;
$X_{11} = X_2 \cdot X_4$;
$X_{12} = X_2 \cdot X_5$;
$X_{13} = X_2 \cdot X_6$;
$X_{14} = X_2 \cdot X_7$;
$X_{15} = X_2 \cdot X_8$; and
$X_{16} = X_2 \cdot X_9$.

Now several models can be defined:

Row Effect:

$$Y = b_0 + b_1X_1 + b_2X_2 + e_1; \quad (1)$$

Column Effect:

$$Y = b_0 + b_1X_1 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + e_2; \quad (2)$$

Row and Column Effects:

$$Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7 + b_8X_8 + b_9X_9 + e_3; \quad (3)$$

and

Full Model:

$$Y = b_0 + b_1X_1 + \ldots + b_{16}X_{16} + e_4, \quad (4)$$

where the b_i are regression coefficients which can differ from equation to equation (i.e., b_1 will not necessarily be unchanged from equation to equation). Table 2 contains sums of squares generated from equations 1-4.

Table 2

Summary Table for Species Data

	df	SS	MS	F
Rows	2	11.18	5.59	.08
Columns	8	157.24	19.66	.28
Rows & Columns	9	158.93		
Rows (Independent of Columns)	1	1.69	1.69	.02
Columns (Independent of Rows)	7	147.75	21.11	.30
Full Model	16	205.82		
Interaction	7	46.89	6.70	.10
Error	187	13113.76	70.13	
Total	203	13319.58		

While in every case the F values are non-significant, perhaps the traditional hypotheses tested are of less interest than investigating the interrelationships (Table 2). If an unadjusted main effects solution (Williams 1972) is preferred, the sources of variation would include rows, columns, interaction and error; the corresponding degrees of freedom are, respectively, 2, 8, 7, 187, or a total of 204 (apparently 1 df more than customary). If a fitting constants solution (Overall and Spiegel 1969) is preferred, the sources of variation are rows (independent of columns), columns (independent of rows), interaction and error; the degrees of freedom are, 1, 7, 7, 187, or a total of 202 (1 df less than usual) respectively. A hierarchical model (Cohen 1968) would of course be additive; however, there would be attendant complications of the hypotheses tested.

The Use of Full Rank Models

Timm and Carlson (1975) gave full rank models a strong impetus for usage with disproportional data; Dalton (1977) recently affirmed their usage. To use

the full rank models, X_1 - X_{16} are slightly modified:

X_1^* = X_1 except $X_1^* = -1$ for scores in the lower samples;

X_2^* = X_2 except $X_2^* = -1$ for scores in the lower samples;

X_3^* through X_9^* are respectively the same as X_3 through X_9 except X_3^* through $X_9^* = -1$ for NW samples;

X_{10}^* through X_{16}^* are product vectors using X_3^* through X_9^* in a manner analogous to X_{10} through X_{16}.

Because there is some ambiguity as to the nature of the models (due to the crest cell being in both a row and a column by itself) six models are defined:

Full Model:

$$Y = b_0 + b_1X_1^* + b_2X_2^* + \ldots + b_{16}X_{16}^* + e_5, \quad (5)$$

Restricted Model 1 - Rows and Interaction:

$$Y = b_0 + b_1X_1^* + b_2X_2^* + b_{10}X_{10}^* + \ldots + b_{16}X_{16}^* + e_6, \quad (6)$$

Restricted Model 2 - Rows and Interaction, crest excluded:

$$Y = b_0 + b_2X_2^* + b_{10}X_{10}^* + \ldots + b_{16}X_{16}^* + e_7, \quad (7)$$

Restricted Model 3 - Columns and Interaction:

$$Y = b_0 + b_1X_1^* + b_3X_3^* + \ldots + b_{16}X_{16}^* + e_8, \quad (8)$$

Restricted Model 4 - Columns and Interaction, crest excluded:

$$Y = b_0 + b_3X_3^* + \ldots + b_{16}X_{16}^* + e_9, \quad (9)$$

Restricted Model 5 - Rows and Columns:

$$Y = b_0 + b_1X_1^* + \ldots + b_9X_9^* + e_{10}. \quad (10)$$

The results from using the full rank models are shown in Table 3.

Table 3

Results from Full Rank Models

Equation	df	$df_{Full} - df_{Rest}$	SS	$SS_{Full} - SS_{Rest}$
5	16	16	205.82	205.82
6	9	7	58.07	147.75
7	8	8	48.58	157.24
8	15	1	204.13	1.69
9	14	2	194.64	11.18
10	9	7	158.93	46.89

Equations 6 and 8 correspond to the results for columns (independent of rows) and rows (independent columns) for the fitting constants method shown in Table 2; also, equations 7 and 9 correspond to the columns and rows effects for the unadjusted main effects solution in Table 2. The interaction in either case is found from the use of equation 10 together with the full model. It can be seen that, for data with the pecularities of the present design, the full rank models do not yield new sets of results as will occur with either disproportional or proportional (but unequal sized cells) data.

Hypotheses Testing through Restricted Models

Using the methodology of Bottenberg and Ward (1963) and Jennings (1967) 17 cell variables are first defined (Table 4).

Table 4

Representation of Cell Model for Mining Cells Data

	Crest	N	NE	E	SE	S	SW	W	NW
Crest	X_1	-	-	-	-	-	-	-	-
Upper	-	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Lower	-	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}	X_{15}	X_{16}	X_{17}

For example, $X_1 = 1$ if the score is from a crest sample, 0 otherwise. All other variables X_2 - X_{17} are analogously defined (note that variables X_1 - X_{17} presented in Table 4 are different from variables X_1 - X_{17} given earlier).

A full model can be defined by

$$Y = b_1X_1 + b_2X_2 + b_3X_3 + \ldots + b_{17}X_{17} + e_{11}, \quad (11)$$

Note that equation 11 does not include a unit vector, implied by the absence of b_0. Likely hypotheses that express the row effect are

$$b_1 = \frac{b_2 + b_3 + b_4 + b_5 + b_6 + b_7 + b_8 + b_9}{8} =$$

$$\frac{b_{10} + b_{11} + b_{12} + b_{13} + b_{14} + b_{15} + b_{16} + b_{17}}{8}.$$

If these restrictions are expressed in terms of two of the coefficients (say b_1 and b_{10}), then

$$b_1 = \frac{b_2 + b_3 + b_4 + b_5 + b_6 + b_7 + b_8 + b_9}{8} \text{ and}$$

$$b_{10} = b_2 + b_3 + b_4 + b_5 + b_6 + b_7 + b_8 + b_9 - b_{11} - b_{12} - b_{13} - b_{14} - b_{15} - b_{16} - b_{17}.$$

If these two restrictions are placed on the full model,

$$Y = \frac{(b_2 + b_3 + b_4 + b_5 + b_6 + b_7 + b_8 + b_9)}{8} X_1 + b_2X_2 + \ldots + b_9X_9$$

$$+ (b_2 + b_3 + b_4 + b_5 + b_6 + b_7 + b_8 + b_9 - b_{11} - b_{12} - b_{13} - b_{14} - b_{15} - b_{16} - b_{17})X_{10} + b_{11}X_{11} + \ldots + b_{17}X_{17} + e_{12}. \quad (12)$$

Rearranging terms,

$$Y = b_2 (1/8X_1 + X_2 + X_{10}) + b_3 (1/8X_1 + X_3 + X_{10}) + b_4 (1/8X_1 + X_4 + X_{10}) + b_5 (1/8X_1 + X_5 + X_{10}) + b_6 (1/8X_1 + X_6 + X_{10}) + b_7 (1/8X_1 + X_7 + X_{10}) + b_8 (1/8X_1 + X_8 + X_{10}) + b_9 (1/8X_1 + X_9 + X_{10}) + b_{11} (X_{11} - X_{10}) + b_{12} (X_{12} - X_{10}) + b_{13} (X_{13} - X_{10}) + b_{14} (X_{14} - X_{10}) + b_{15} (X_{15} - X_{10}) + b_{16} (X_{16} - X_{10}) + b_{17} (X_{17} - X_{10}) + e_{12}. \quad (13)$$

Now, 15 variables are defined in relationship to b_2 through b_9 and b_{11} through b_{17}. For example, the variable related to b_2 is $V_2 = 1/8$ if from a crest sample, 1 if from a cell 2 sample, 1 if from a cell 10 sample, 0 otherwise. The restricted model, after creating all 15 variables would be

$$Y = b_2V_2 + b_3V_3 + \ldots + b_9V_9 + b_{11}V_{11} + \ldots + b_{17}V_{17} + e_{12}. \quad (14)$$

Programs such as those of McNeil et al. (1975) LINEAR or Ward and Jennings (1973) DATRAN can use equations like (14) above directly. Most multipurpose multiple linear regression programs automatically include a unit vector, thus including the b_0 term. Any *one* of the variables V_2 - V_9 or V_{11} - V_{17} can be dropped if a multipurpose multiple regression program is used; for the present purpose V_{17} is (arbitrarily) dropped. The restricted model becomes

$$Y = b_0 + b_1V_1 + b_2V_2 + \ldots + b_9V_9 + b_{11}V_{11} + \ldots + b_{16}V_{16} + e_{12}. \quad (15)$$

Using equation 15 as the restricted model, df = 14, SS = 194.64; thus, the df for rows is 2, and SS_{rows} 205.82 - 194.64 = 11.18, identical to those values found both with the use of equation 1 and with the difference between equations 5 and 9.

A similar solution could be set for columns as well. A set of hypotheses likely to be of interest for the column effect is (in terms of the full model, equation 11)

$$b_1 = \frac{b_2 + b_{10}}{2} = \frac{b_3 + b_{11}}{2} = \frac{b_4 + b_{12}}{2} = \frac{b_5 + b_{13}}{2} = \frac{b_7 + b_{15}}{2} = \frac{b_8 + b_{16}}{2} = \frac{b_9 + b_{17}}{2}.$$

In a manner similar to that shown for the row effect a restricted model can be found such that df = 8, SS = 48.58, yielding for the columns effect df = 8, $SS_{columns}$ = 157.24, results similar to those found for equation 2 and for the difference between equations 5 and 7.

Thus, while it may be somewhat troublesome initially to find that the degrees of freedom, if totaled, exceed the total degrees of freedom (by one), this oddity arises from testing what would appear to be a useful set of hypotheses. Except for hierarchical models, those who have used disproportionate cell frequencies are already familar that such data are generally non-additive with respect to both the sums of squares and R^2 values; in some situations with missing cells the degrees of freedom may become non-additive.

The Effect of Dropping a Vector on the Hypotheses Tested

Some readers, particularly those familiar with the paper by Jennings (1978), might wonder about the effect of dropping a vector, such as V_{17} in equation 15.

Strictly speaking, the dropping of the V_{17} from the full model (equation 11) should be written

$$Y = b_0 + c_1X_1 + c_2X_2 + c_3X_3 + \ldots + c_{16}X_{16} + e_{13} \quad (16)$$

where each $c_i = b_i - b_0$.

Also, $b_0 = \bar{Y}_{17}$; thus each $c_i = \bar{Y}_i - \bar{Y}_{17}$ for equation 16. Letting $b_{17} = 0$ in equation 11 is called setting a side condition; equation 16 is a re-parameterization of equation 11 (meaning that the expected values for all Y_i are the same in each equation; also, $e_{11} = e_{13}$). The null hypothesis corresponding to the row effect would be

$$c_1 = \frac{c_2 + \ldots + c_9}{8} = \frac{c_{10} + \ldots + c_{16} + b_0}{8}.$$

It should be noted that this expression of the row effect is testing the same hypothesis as stated earlier; if b_0 is added to each portion of the present expression, the previous expression for the row effect would occur. The restrictions become (in terms of c_1 and c_{10})

$$c_1 = \frac{c_2 + \ldots + c_9}{8} \text{ and } c_{10} = c_2 + \ldots + c_9 - (c_{11} + \ldots + c_{16} + b_0).$$

Placing these restrictions on equation 16,

$$Y = \frac{b_0 + (c_2 + \ldots + c_9)}{8} X_1 + c_2X_2 + \ldots + c_9X_9 + [(c_2 + \ldots + c_9) - (c_{11} + \ldots + c_{16} + b_0)]X_{10} + c_{11}X_{11} + \ldots + c_{16}X_{16} + e_{14}. \quad (17)$$

Rearranging terms,

$$Y = b_0 + c_2(1/8X_1 + X_2 + X_{10}) + c_3(1/8X_1 + X_3 + X_{10}) + c_4(1/8X_1 + X_4 + X_{10}) + c_5(1/8X_1 + X_5 + X_{10}) + c_6(1/8X_1 X_6 + X_{10}) + c_7(1/8X_1 + X_7 + X_{10}) + c_8(1/8X_1 + X_6 + X_{10}) + c_9(1/8X_1 X_7 + X_{10}) + c_{11}(X_{11} - X_{10}) + c_{12}(X_{12} - X_{10}) + c_{13}(X_{13} - X_{10}) + c_{14}(X_{14} - X_{10}) + c_{15}(X_{15} - X_{10}) + c_{16}(X_{16} - X_{10}) + e_{14}. \quad (18)$$

The similarity between equations 13 and 18 is striking; also $e_{12} = e_{14}$.

An equation similar to equation 14 can be constructed:

$$Y = b_0 + c_2V_2 + c_3V_3 + \ldots + c_9V_9 + c_{11}V_{11} + \ldots + c_{16}V_{16} + e_{14}. \quad (19)$$

Actually, equation 19 is the analog to equation 15; equation 15 would more correctly be said to be the same as equation 19. That is, the values for the b_i in equation 15 are actually the c_i from equation 19. Similar models could be developed for the columns as well.

The hypotheses tested by equation 19 use regression coefficients rather than the means. The hypothesis in terms of the population means corresponding to the restriction on equation 11 (not including the unit vector) is

$$\mu_1 = \frac{\mu_2+\mu_3+\ldots+\mu_9}{8} = \frac{\mu_{10}+\mu_{11}+\ldots+\mu_{17}}{8}. \qquad (20)$$

The restriction on equation 16 (including the unit vector) corresponds to the following hypothesis regarding the population means:

$$(\mu_1-\mu_{17}) = \frac{(\mu_2-\mu_{17})+(\mu_3-\mu_{17})+\ldots+(\mu_9-\mu_{17})}{8} =$$

$$\frac{(\mu_{10}-\mu_{17})+(\mu_{11}-\mu_{17})+\ldots+(\mu_{16}-\mu_{17})}{8} \qquad (21)$$

If μ_{17} is added to each side of the last expression, equation 20 would result.

Other Applications With The Present Design

The present design would be applicable whenever double entry tables are used wherein all sampling units except for a control group are placed into a treatment combination that includes two actual treatments. The present design allows a control group within the context of a two-way analysis of variance; other designs may include a no-treatment for each dimension of the design. If there is an intrinsic interaction, the interaction might be overlooked in designs that include the no-treatment group for both classifications crossed with other treatments in the two-way classification.

As an example of such an experiment in reclaiming strip-mined land, suppose an amendment program began so that four (non-zero) levels of potassium are used as one way of classification together with the second way of classification consisting of several amendments. If the interest is in looking at possible interactions together with looking for the most effective treatment combination, then perhaps a single control group would serve the experiment as well as trying to incorporate a control group for each way of classification by crossing them with the treatments of the other dimension.

A very different sort of biological application could be made in the area of genetics. Suppose there is interest in cross-breeding an available strain of laboratory mice with three genetically different strains. The intent is to measure hybrid vigor in the offspring when compared to similar offspring in the available colony; further, care is taken to match animals from the four strains such that sex of the parent is also a factor. A table similar to Table 1 could be constructed so that the experiment could be executed:

		1	2	3
Orig.	X_1	-	-	-
Male	-	X_2	X_3	X_4
Female	-	X_5	X_6	X_7

Except for cell 1 (X_1), all other cells would contain hybrids. It would make sense to view cell 1 as a control group to measure the other cells against.

In ecological studies, consider an area such as that shown in the following diagram, in which a natural transition occurs from a grassland to forest. Assume that the burned and mowed/clearcut areas were well defined and were established at the same time. Also assume that a relatively long period of time had passed between establishment and sampling. The objective of the experiment would be to determine the similarities (1) in species composition by some appropriate measurement between the untreated grassland and each of the treatments (burned former grassland, burned former forest, mowed former grassland, clearcut former forest), and (2) the similarities of the untreated forest to each treatment as well as the effects of burning vs. mowing, former forest vs. former grassland.

2

cut 4 6 Burned area

Forest transition

Grassland

3 5

1

The table could then be depicted as

	No treatment	Cut	Burned	No treatment
Grassland	X_1	-	-	-
Former grassland	-	X_3	X_5	-
Former forest	-	X_4	X_6	-
Forest	-	-	-	X_2

Educational Applications

While the example given was concerned with sampling plant species in strip-mined land and the other applications had a biological orientation as well, one might ask, "What value is the design in human experimentation?" Within a design sense, an experimenter may have a two-way treatment classification procedure; however, a true control group would be administered neither treatment combination. In such a case, the crest cell would correspond to the control group and the remaining rows and columns would constitute the treatments.

Yet another application might arise from a concern for the cultural bias associated with a given test. Suppose, for sake of argument that a critique of the Iowa Test of Basic Skills (ITBS) asserts that, in regard to district wide means, all the ITBS measures is the distance from Iowa (the further the distance, the lower the mean score). Adjacent states might constitute the "close" states and the next layer of states might constitute the "farther" layer of states (undoubtedly a much more sophisticated sampling process would occur with the actual choosing of subjects). In diagramatic form, the cells would be represented as

	Inner	North	N East	East	South	West	N West
Inner	Iowa	-	-	-	-	-	
Close	-	Minnesota	Wisconsin	Illinois	Missouri	Nebraska	S Dakota
Farther	-	N Dakota	Michigan	Indiana	Arkansas	Wyoming	Montana

Random samples of school districts could be chosen from the listed states and the original hypotheses could be tested.

REFERENCES

Bottenberg, R.A. and Ward, J.H. *Applied multiple linear regression*. Lackland Air Force Base: Personnel Research Laboratory, TDR-63-6, 1963.

Cohen, J. Multiple regression as a general data-analytic system. *Psychological Bulletin*, 1968, 70, 426-443.

Dalton, S. Regression approaches and approximate solutions to analysis of variance with disproportionality varied. *Multiple Linear Regression Viewpoints*, 1977, 7, No. 2, 16-32.

Jennings, E. Fixed effects analysis of variance by regression analysis. *Multivariate Behavioral Research*, 1967, 2, 95-108.

Jennings, E. Fixed effects analysis of variance with unequal cell sizes. *Journal of Experimental Education*, 1978, 46, No. 2, 42-51.

McNeil, K.A., Kelly, F.J. and McNeil, J.T. *Testing research hypotheses using multiple linear regression*. Carbondale, Ill.: Southern Ill., Univ. Press, 1975.

Overall, J.E. and Spiegel, D.K. Concerning least squares analysis of experimental data. *Psychological Bulletin*, 1969, 72, 311-322.

Searle, S.R. *Linear models*. New York: Wiley, 1971.

Timm, N.J. and Carlson, J.R. Analysis of variance through full rank models. *Multivariate Behavioral Research: Monograph*, 1975, 75-1.

Wali, M.K. *A systems approach to the reclamation of strip-mined areas in North Dakota*. Grand Forks: Project Reclamation, Univ. North Dakota, 1977.

Ward, J.H. and Jennings, E. *Introduction to linear models*. Englewood Cliffs, N.J.: Prentice-Hall, 1973.

Williams, J.D. Two way fixed effects analysis of variance with disproportionate cell frequencies. *Multivariate Behavioral Research*, 1972, 7, 67-83.